



# Logistic Regression

Week 4

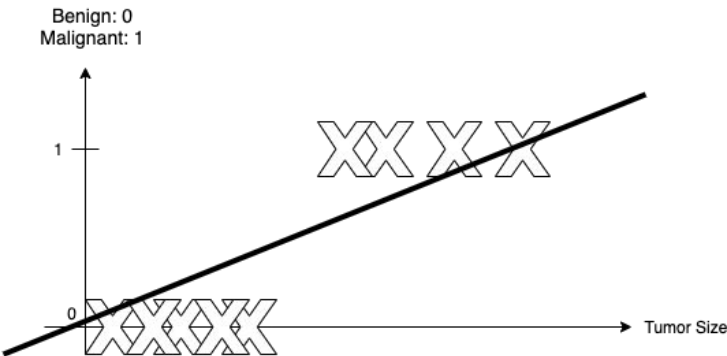
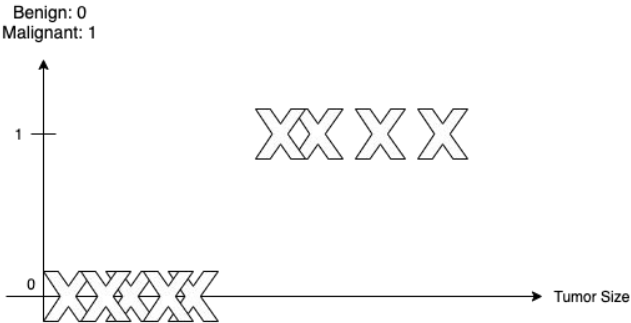


# Linear Regression for Classification

We have a dataset that is represented as below

We want to fit a function that is able to capture the relation, and fit best as possible

Using linear regression we will get such result;





# Problems

Not perfectly fit the data.

Inaccurate results, we want our results to be  $\{0, 1\}$  but Linear Regression will predict more than 1 which is nonsense.

We need more suitable function that is able to fit the data better than Linear Regression

# Logistic Regression





# Forward Pass

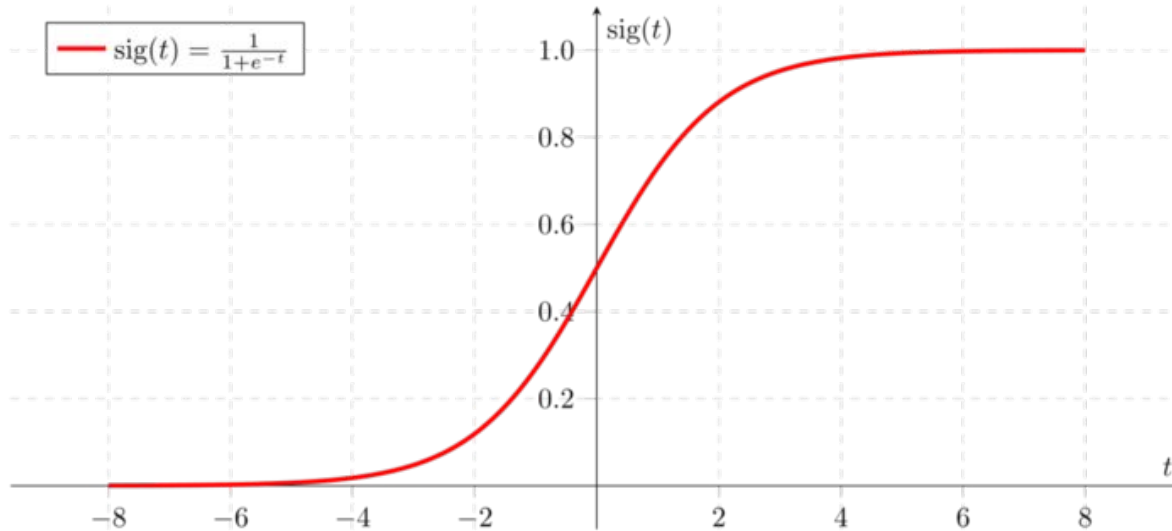
Same in Linear Regression, we have an polynomial equation which multiplies the weights and inputs.

$$y = mX + b$$

Additionally, we should use another function which will fit the previous data perfectly.

That is one and only **Sigmoid Function**

# Sigmoid Function

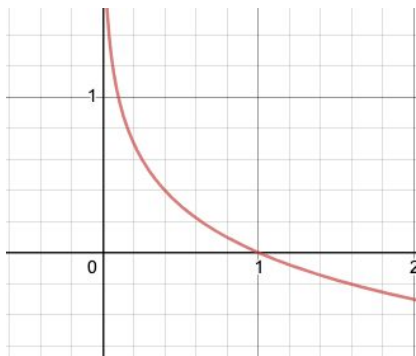


# Loss Function

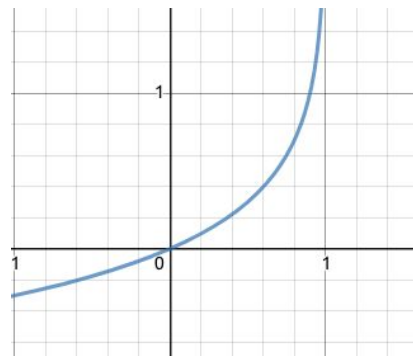
$$J(w) = \frac{1}{m} \sum_{i=1}^m \text{Cost}(f(x_i), y_i)$$

$$\text{Cost}(w) = -y(\log(f_w(X))) - (1 - y)(\log(1 - f_w(X)))$$

$y = 1$



$y = 0$



Check the full derivation [here](#)



# Optimization

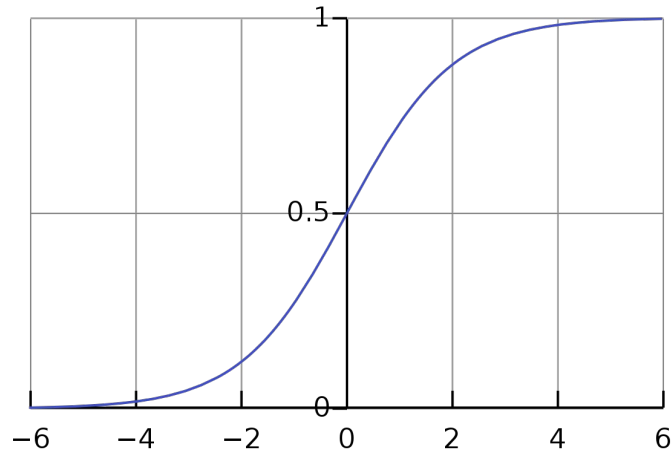
$$\sigma'(z) = \sigma(z)(1 - \sigma(z))$$

$$\frac{\partial J}{\partial w_j} = \frac{-1}{m} \sum_{i=1}^m (y^i - f_w(x^i))(x_j^i)$$

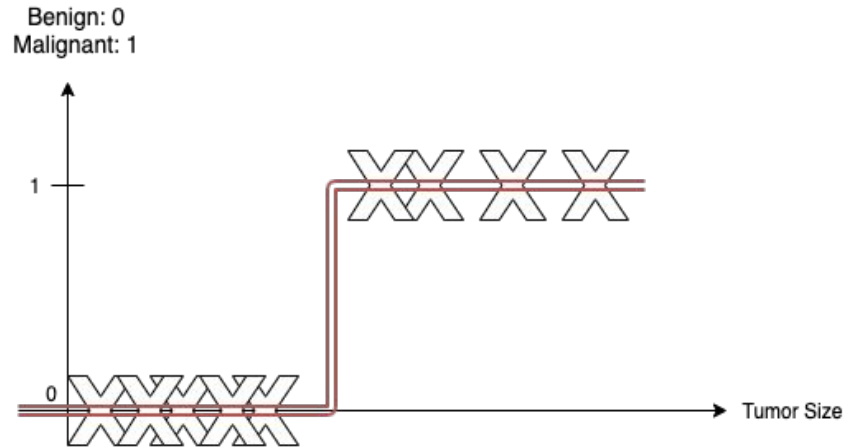
$$w_j := w - \alpha * \frac{\partial J}{\partial w_j}$$



# Making Prediction



$$Decision = \begin{cases} 1, & P(class = 1|X) \geq 0.5 \quad \text{or} \quad \sigma(z) \geq 0.5. \\ 0, & P(class = 1|X) < 0.5 \quad \text{or} \quad \sigma(z) < 0.5. \end{cases}$$





# Today's Project

Sentiment Analysis of sentences from 2 different company feedback forms (Yelp and Amazon)

Firstly, we will preprocess words into their stems

Secondly, we will count the number of words pass in positive and negative sentences

After extracting the frequency table, we will prepare our features and train our model in **PyTorch**



# Simple Example

Sentence 1: "I enjoyed the ride"

Sentence 2: "I hated the ride"

Sentence 1 score: [4, 3]

Sentence 2 score: [3, 4]

Words	Positive #	Negative #
I	1	1
Enjoy	1	0
The	1	1
Ride	1	1
Hate	0	1